AD-A242 213

Lehigh University

telephone (215) 758-3950

Sherman Fairchild Center for Solid State Studies 161
Bethlehem, Pennsylvania 18015-3185

DTIC
ELECTE
NOV1 1991
S D
C

OFFICE OF NAVAL RESEARCH

SEMIANNUAL TECHNICAL REPORT

for

1 March 1991 through 1 September 1991

for

Contract N00014-89-J-3149

R&T PROJECT: 1d14001--01

Title of Contract

"Electrically Modifiable Nonvolatile SONOS Synapses for Electronic Neural Networks"

Name of Principal Investigator

Dr. Marvin H. White

Name of Organization

Lehigh University
Sherman Fairchild Center - Bldg. 161
Bethlehem, PA 18015
(215) 758-4421

Dr. Clifford Lau
Department of the Navy
Office of the Chief of Naval Research
800 North Quincy Street, Code 1114SE
Arlington, VA 22217-5000
(202) 696-4961

91-13592

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

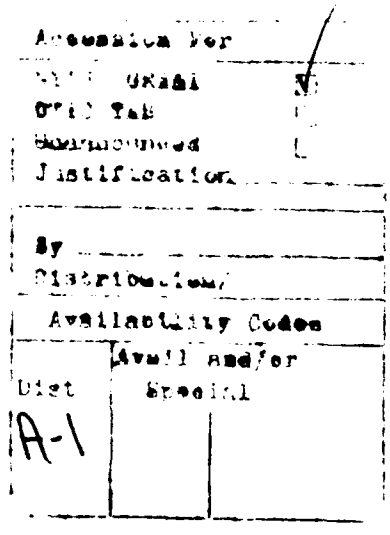
Semi-Annual Report for

DARPA/ONR

by

Marvin H. White, Chun-Yu Malcolm Chen, and Margaret French

Lehigh University
Sherman Fairchild Center
Bethlehem, PA 18015



Statement A per telecon
Clifford Lau ONR/Code 1114
Arlington, VA, 22217-5000

NWW 10/31/91

Abstract

This research addresses the implementation of an electronic element, which emulates the biological synaptic interconnection, in an artificial electronic neural system. The basic interconnection, or the weight, consists of an electrically reprogrammable, nonvolatile, analog conductance which programs at 5V levels. In addition, the fabrication technology for this synaptic interconnection is compatible with existing CMOS VLSI process. The attractive features of this synaptic weight will be discussed in this report. Furthermore, this report examines the material needs, the device structures, the use of the synaptic weights in a two-tap weight linear adaptive neural-like circuit and the issue of integrating both the synaptic weight elements and the peripheral circuit onto a single silicon wafer.

1. Introduction

The current surge of enthusiasm for neural network aims to construct systems that can learn or modify their behavior according to the environment. There are many similarities which exist between this new class of machine and human beings. One of these similarities is the massive parallelism in processing information. Parallel processing[1] concepts are in stark contrast to the operations of modern digital computers that perform large numbers of sequential operations very rapidly and accurately.

Researchers believe the **synaptic junctions** in a neural system are the local memory sites and provide the physiological basis for the distributed parallel systems.[2, 3] These synapses are not only modifiable but also serve the functions of storing and transmitting information from neuron to neuron. To reduce the complex modelling required for the synaptic interconnection, the representation of the synapse has been simplified to a single ideal junction between the output of neurons (axons) and the inputs to neurons (dendrites). Synaptic modification requires information from the input and the output of the neuron in order to perform complex recognition. Therefore, the nature of the synaptic junction and the principle or algorithm which controls local organization at the neuron level become two central issues pertaining to neural networks research.

The recent interest in neural networks[4, 5] is a direct consequence of the programmability which is an essential feature of learning machines, associative memories, and adaptive signal processors. Programmability requires a *modification of the synaptic strength* in the language of neurobiology. If we seek an efficient hardware implementation of electronic neural systems, then the synapses - as well as the network itself- should be analog. Several attempts have been made to realize programmable synapses, either digitally[6] or with temporary storage on the input capacitance

of a MOS Transistor[7,8] to alter the latter's analog conductance. The former approach stores the weight information in digital registers and thus suffers from excessive chip area and power consumption. On the other hand, although the MOS Transistor provides an analog synaptic strength (weight) in a small chip area, the weight is temporary and requires periodic refresh similar to a DRAM. Thus, this dynamic refresh approach lacks the nonvolatility and storage properties of an EEPROM cell. Researchers at Intel have reported an electrically trainable artificial neural network with floating gate device as the synaptic element.[9] Although floating gate device has the property of nonvolatility, its high programming voltage requirement prevents it from being technologically compatible with scaled CMOS process.

In this research report we describe a new approach to obtain an electrically reprogrammable or modifiable synaptic weight to be used as a basic functional element in electronic neural systems. The salient features of this network element are the following:

- **Low programming voltages(5-10V) which are compatible with peripheral CMOS VLSI technology in contrast with Floating Gate approaches.**
- **Low power dissipation ($< 1\mu W$).**
- **Dynamic Range of 1000:1 (60 dB).**
- **Nonvolatile features which mimic biological synapses with respect to memory loss (e.g. 20% of the information available after 10 years) and reinforced learning (e.g. successive interrogation enhances memory retention).**
- **Small synaptic area on a VLSI chip (e.g. less then $20\mu m^2$ for 1.25 μm feature sizes).**
- **Extensive erase/write programming cycles are possible with this synapse ($> 10^8$ cycles) in contrast with Floating Gate approaches.**
- **Inherent radiation damage resistance beyond a total dosage of 1MRad (Co^{60}) and 10^9 Rad/sec transient which is not possible with Floating Gate technology. Thus, if radiation damage resistance of neural networks is an important issue. then the SONOS devices have demonstrated success in this area.**

The basic nonvolatile device structure, which we describe in this report was first introduced as a digital nonvolatile memory cell in the summer of 1987 at the IEEE Device Research Conference[10] by researchers at Lehigh University. We have had a continual involvement over a 20 year period with nonvolatile memories, beginning in the late 60's where we had programming voltages of 25V. to the late 80's with our novel 5V SONOS device structures. During this time period we introduced the use of CCD's and nonvolatile memories[11,12,13] in nonvolatile charge addressed memories (NOVCAM). These ideas have been employed recently for neural network circuits by researchers at

Lincoln Laboratories.[14] Our recent work recognizes the inherent analog conductance aspect of the nonvolatile SONOS memory device which makes it a perfect candidate for the modifiable synapse in an electronic neural system.

In addition to the realization of an electronic element to simulate the synaptic interconnections of a neural network, we must have a method or algorithm to change or reprogram these interconnections and, thus, alter the connectivity of the neural network. We have had experience with a particular form of an algorithm, namely, the Widrow-Hoff Least Mean Square (LMS)[15] error algorithm or in neural network terminology - the so-called 'delta rule'. In the late 70's we researched a CCD Adaptive Analog Signal Processor[16, 17] which realizes the 'delta rule' with CCD analog delay lines and electrically reprogrammable MNOS analog conductance weights. These weights were nonvolatile memory transistors whose analog conductance was programmed with voltages ranging from 15-25V. Our recent work on 'scaling' these programmable analog conductances has resulted in a new device structure, called the SONOS nonvolatile memory transistor, which can be reprogrammed with voltages ranging from 5-10V. This work has recently been described at the 1991 11th IEEE Nonvolatile Semiconductor Memory Workshop.[18] These voltage levels are compatible with 'scaled' CMOS VLSI technology which has 12-15V breakdown voltages for 1.25μm feature sizes. In this report we describe our recent work on the electrically reprogrammable (modifiable) SONOS nonvolatile synapse and a simple electronic neuron with 2 synaptic weights. We discuss this two-tap weight linear adaptive neuron in terms of the technology, the electrical characteristics of the synapses, and their performance in this simple test vehicle - a 'delta rule' adaptive signal processor.

2. Technology and Characterization of SONOS Synaptic Weight

The programmable synapse is the result of an ongoing effort at Lehigh University to 'scale' the programming voltages required to alter the analog conductance of a nonvolatile memory transistor with a multi-layer (oxide-nitride-oxide) gate insulator as shown in Fig. 1. Recent efforts in scaling this device have resulted in a SONOS (Silicon/Blocking Oxide/Nitride/Tunneling Oxide/Silicon) nonvolatile memory transistor which is electrically reprogrammable at CMOS voltage levels. Typically, the tunneling oxide is 15-25Å, the storage nitride is 50-100Å and the blocking oxide is 35-50Å. Fig.2 shows the Transmission Electron Microscope (TEM) photograph of the cross sectional view of the SONOS transistor. This device is similar to a SNOS transistor except for the addition of the blocking oxide which is used to inhibit injection of carriers from the polysilicon gate electrode

and also to improve the memory retention by prohibiting the transfer of stored charge from the nitride to the gate electrode. As the result, the blocking oxide permits the entire dielectric sandwich to be scaled to dimensions where the programming voltages ranging from 5-10 V are possible.

When the SONOS device is subjected to a positive (or negative) programming pulse, electrons (or holes) are injected into the silicon nitride layer by means of tunneling across the thin tunnel oxide. The injected charges are trapped by the silicon nitride and thus shift the threshold voltage positively (or negatively). The threshold voltage of a SONOS transistor can be written as

$$V_{TH} = \phi_{GS} - \frac{Q_f}{C_{eff}} + \left(\frac{x_{ob}}{\varepsilon_{ox}} + \frac{x_n - \bar{x}}{\varepsilon_N} \right) Q_N + 2\phi_B + \frac{\sqrt{4\,\varepsilon_{si}\, q\, N_B\, \phi_B}}{C_{eff}} \qquad (1)$$

where ϕ_B is the bulk potential, ϕ_{GS} is the gate to semiconductor workfunction, Q_f is the fixed charge at the tunneling oxide-silicon interface, ε_{ox} and ε_N are the dielectric permittivities of the oxide and nitride, ε_{si} is the dielectric permittivity of the bulk silicon, x_{ot} is the tunnel oxide thickness, x_{ob} is the blocking oxide thickness, x_n is the nitride thickness, $\bar{x}$ is the charge centroid in the insulator, and Q_N is the charge stored in the nitride, N_B is the bulk doping density, and

$$C_{eff} = \frac{\varepsilon_{ox}}{x_{ot} + \frac{\varepsilon_{ox}}{\varepsilon_N} x_n + x_{ob}} \qquad (2)$$

We assume that the tunnel oxide and blocking oxide have the same dielectric permittivity, even though, it is known that the tunnel oxide is silicon rich and the blocking oxide is an oxynitride. The values of the charge centroid $\bar{x}$ and the charge stored in the nitride Q_N will change as the device is written or erased. The analog conductance of the SONOS synaptic weight may be written as

$$g_{ds} = \bar{\mu}_{eff} \frac{W}{L} C_{eff} (V_{GS} - V_{TH}) \qquad (3)$$

where $\bar{\mu}_{eff}$ is the effective carrier mobility, V_{GS} is the read voltage, and V_{TH} is the electrically modifiable threshold voltage given in equation (1). Therefore, there are two ways which the analog channel conductance can be altered: (1) change the value of V_{GS} or (2) change the value of V_{TH}. In our study, the latter approach is chosen.

The SONOS transistors are characterized for their memory properties by using the test station described in Anirban Roy's Master's Thesis.[19] This test station allows one to take both the

erase/write and retention measurements. To investigate the memory loss/retention properties of the synaptic weight element, retention measurements are taken. The retention characteristics are obtained by applying positive (negative) five volts to the gate for 10 seconds to place the device in the write (erase) state and then measuring the turn-on voltage after a varying delay time. The turn-on voltage is related to the threshold voltage by

$$V_T = V_{TH} + \sqrt{\frac{2I_{DS}}{\beta}} \tag{4}$$

with I_{DS} as the forced drain to source current during measurement and

$$\beta = \bar{\mu}_{eff} \left(\frac{W}{L}\right) C_{eff} \tag{5}$$

where W is the width of the transistor, L is the length of the transistor, and $\bar{\mu}_{eff}$ is the effective mobility. The effective mobility is the bulk mobility reduced by Coulombic and surface scattering of carriers in the inversion layer. This mobility is influenced by the gate and substrate voltages.[20] For a SONOS transistor retention measurements indicate that greater than 20 percent of the memory window remains after a projected 10 year delay time as shown in Fig. 3. The erase/write measurements indicate the programming speed of the synaptic weight element. To measure the writing (erasing) speed, negative (positive) five volts are applied to the gate for 10 seconds to place the device in the erase (write) state. Then, positive (negative) five volts are applied to the gate with varying pulse widths and the turn-on voltage is measured after each pulse width. The erase/write characteristics of the SONOS memory transistor are shown in Fig. 4. A wide dynamic range is one of the essential properties for the synaptic weight element, and Fig. 5 illustrates a 60 dB in dynamic range after ± 5V programming for the SONOS synaptic weight. In addition, a recent study in reliability has demonstrated the inherent resistance of the SONOS memory transistor to radiation damage ($\delta V_{TH} = 0.1$V, with $V_{GS} = +5$V at 1MRad Co^{60} radiation).[21]

3. Single-level Linear Adaptive Neuron

We have incorporated the SONOS synaptic weights into a single-level linear neuron-like circuit using a Widrow-Hoff's delta learning rule.[15] The circuit is built with a hybrid breadboard of CMOS components for the control logic and the algorithm implementation and the SONOS nonvolatile memory transistors to demonstrate the voltage level compatibility of both SONOS and CMOS technologies. Many researchers believe that the neural system is made up of several layers

of neurons and Fig. 6 shows the multi-layer architecture of an artificial neural network. The first layer of neurons, the input layer, can be best thought as the sensory neurons in a human body. The weight connections between the input layer and the middle hidden layer are normally considered to be feedforward and fixed. On the other hand, the weight connections between the middle hidden layer and the output layer are considered to be feedback in nature. Our work has concentrated on the implementation of two neurons in the hidden layer and one output neuron as highlighted in the figure.

Fig. 7 shows the block diagram of the single-level linear adaptive neuron. A desired response (or external teacher), $d(m)$, is presented to the neuron as the training signal. If the output of the linear adaptive neuron is not trained, then there exists a mismatch between the output of the linear adaptive neuron, $y(m)$, and the desired response, $d(m)$.

$$\varepsilon(m) = d(m) - y(m) \tag{6}$$

where $\varepsilon(m)$ is the error generated. This error is then used by a learning algorithm, namely the Clipped-data Least Mean Error algorithm, to minimize the error generated and thereby training the neuron to the correct response. This single-level linear adaptive neuron has two tap weights, each weight is composed of two SONOS analog electrically reprogrammable conductances as shown in Fig. 8. Since the synaptic weight must be either positive or negative in value, we have chosen a differential weighting scheme. If the analog conductance connecting the positive summing path to the differential operational amplifier is greater than the analog conductance connecting the negative summing path to the differential operational amplifier, then the weight is positive in value. On the other hand, if the opposite case is true, then the weight is negative in value. Positive weight value corresponds to the excitatory synaptic strength and the negative weight value corresponds to the inhibitory synaptic strength.

In operation, the input signal $x(n)$ is passed through a switched capacitor analog delay line where the input signal is sampled and delayed to create two tapped signal outputs $x_0(m)$ and $x_1(m)$. These tapped signals multiply to their corresponding programmable weights W_0 and W_1 and the result is summed linearly at the summing amplifier. The output $y(m)$ can be expressed as:

$$y(m) = \sum_{i=0}^{1} W_i(m) \cdot x_{m-i} \tag{7}$$

where m is the time index and k is the spatial index. A correlated double sampling technique[22] is employed in the circuit to remove the unwanted noise and offset voltages introduced by the summing amplifiers. The linear adaptive neuron is configured to perform Widrow-Hoff's delta rule as:

$$W_k(m+1) = W_k(m) + \Delta W_k(m) \tag{8}$$

where $\Delta W(m)$ is the incremental weight to be calculated by the clipped-data least mean square error (C-LMSE) algorithm[23]:

$$\Delta W_k(m) = 2\mu\,|\varepsilon(m)| \cdot Sgn[\varepsilon(m)]\,Sgn[x(m-k)] \tag{9}$$

where μ is the convergence factor. Compared to the regular Least Mean Square Error algorithm, the input signal amplitude is clipped in the learning algorithm. This algorithm eliminates the usage of a four quadrant multiplier needed for the LMS error algorithm. The sign multiplication in the incremental weight calculation is essentially an Exclusive OR operation and the output of the Exclusive OR gate controls the path of proper gate programming voltage for the SONOS synaptic weight. If the convergence factor is small, then the system will minimize the misadjustment caused by the variance of the weights; however, this also results in a long convergence time. Conversely, if we choose to use a larger convergence factor, then the convergence time of the system is shortened with the penalty of larger misadjustment. The backpropagating error is used to calculate the adjustments to minimize the system error as shown in equation (9). Once the error is minimized, the system is said to be in its steady state condition[24] where the output of the system, $y(m)$, is the best match of the training signal, $d(m)$, or the 'external teacher'.

The incremental weight update is essentially a cross correlation between the error and the clipped input data vectors. The update stops when the two vectors become orthogonal. Sometimes, the network may be overcorrected initially, however, the error will be quickly minimized by the learning algorithm and the system reaches its desired response. The digital delay line provides the sign information of the input to the learning algorithm. A special steering network is designed to switch the proper programming voltages to the gate terminals of the SONOS transistors once the incremental weights are calculated.

4. Experimental Results

There are two main types of characteristics from which the electrical performance of the linear adaptive neuron can be evaluated. The first characteristic, namely the output and training signals versus time characteristics, gives the information on how well the output signal approximates the training signal especially in the phase relationship between these two signals. The second characteristic, namely the error signal versus time characteristics, shows how fast the linear adaptive neuron adapts before it reaches its minimum error. A typical output and training signals versus time characteristic consists of two parts: the initialized and the adapted part. In the initialized part, the weights are first initialized to a known state (either the fully positive or the fully negative state) and then the weights are subjected to a reading voltage to read out the weight information and the output signal and the training signal are compared and recorded. The linear adaptive neuron is then allowed to adapt itself to the training signal and the results are shown in the adapted part of the characteristics. Figure 9 shows the output and training signal versus time characteristic.

A typical error signal versus time characteristic is obtained with initialized weight values and monitoring the error signal with time. Our observation indicates the weight initialization scheme affects the convergence behavior of the linear adaptive neuron. This phenomenon is attributed to the nonsymmetric *erase* and *write* characteristics of the SONOS transistor. Therefore, one weight initialization scheme may require more erase action taking place than another weight initialization scheme, causing a difference in convergence characteristics. Figures 10 shows a typical error versus time characteristic.

5. Technical Progress

The performance of the linear adaptive neuron depends strongly on the programming characteristics of the synaptic weight elements. Therefore, a fabrication run aimed specifically at improving the programming speed of the synaptic weight elements has been completed and fully characterized. The main differences between the new and the old devices are the thicknesses of the storage nitride layer and the blocking oxide layer. By reducing the nitride layer and increasing the blocking oxide layer thicknesses, the gate injection of carriers is minimized. Furthermore, the charge tunneled across the tunneling oxide is better retained in the nitride layer with a thicker blocking oxide. As a result, the programming speed, indicated by the erase/write characterization of the new devices, demonstrates a roughly one order of magnitude improvement in the programming

speed (determined by the cross-over time as mentioned in the previous section) over the old devices as shown in figure 11. In addition, the programming voltage dependence on the programming speed, as shown in figure 11, illustrates an improvement of one order of magnitude in programming speed with each one volt increment in the programming voltage. We have also incorporated the new synaptic weight elements into the linear adaptive neuron. Since the programming speed of the new synaptic weight elements has improved by a factor of 10, the convergence time (a measure of the performance of the linear adaptive neuron) has been reduced by a factor of 10 accordingly as shown in figure 12.

A fully computer controlled data acquisition system is an invaluable tool for synaptic weight element characterization. The current measurement system requires the operator to manually set up the measurement sequence and hand-record the data obtained. An automatic data acquisition system enables the user to set up the measurements, and analyze the data, all under the control of one console. The block diagram of such a system is shown in figure 13. A vital part of the proposed system, the HPIB command/data interpretor has been successfully designed, constructed, and fully tested for functionality. The schematic of the HPIB command/data interpretor is shown in figure 14.

We have continued our efforts on the integration of the linear adaptive neuron into a single silicon wafer. We have layed out some of the vital parts of the linear adaptive neuron under the Mentor Graphics package with a technology specifically geared to the fabrication process of our microelectronics laboratory at Lehigh. Since we are creating analog ASICs, area and power consumption of the designs have to be minimized. Upon completion of layout, an extraction of the netlist and the parasitics are obtained and SPICE circuit simulations are performed to ensure the designs are properly transferred into layouts. Figures 15 and 16 illustrate a sample of the layouts already completed.

The research efforts during the period of March 1991 to September 1991 have resulted in a paper to be presented at the International Joint Conference on Neural Networks (IJCNN 91) to be held in November at Singapore.

6. Proposed Investigations

We have demonstrated the direct relationship between the SONOS synaptic weight element characteristics and the performance of the linear adaptive neuron. Therefore, we feel it is important to continue our research efforts on the synaptic weight elements. In particular, the investigation of how to make even faster devices has resulted in the fabrication plans of utilizing thinner tunneling oxide thickness and p+ polysilicon gate material.

We have also pledged our effort to the development of the automatic data acquisition system to further aid our research. A customized computer program will be written to drive and interface with the existing measurement equipment. The data obtained can then be collected, formatted, and analyzed with device parameter extraction routines.

We will focus on the layout design of the remaining parts of the linear adaptive neuron. We believe the advent in SONOS device technology and the integration of the signal processing circuitry as well as the synaptic weight element onto a single silicon wafer will make an impact to the Artificial Intelligence Neural Network Technology field.

7. Conclusions

The SONOS nonvolatile memory transistor has been shown to be an ideal electronic element for the electrically reprogrammable analog conductance in an artificial neural network. We have demonstrated the attractive features of this synaptic weight for the use of large neural network systems, for instance, low programming voltage (5-10V), low power dissipation(<1μW / synapse), small chip area (estimated $20\mu m^2$/ weight cell for a 1.2 μm feature size), a dynamic range of 60 dB, good memory retention (20 % window at a projected 10 years period), and endurance beyond 10^7 erase/write cycles. In addition, the SONOS synaptic weight has inherent resistance to radiation damage (ΔV_{th}=0.1V, with V_{gs}=+5V at 1MRad Co^{60} radiation). We have been continuing our efforts in optimizing the modifiable synaptic weights to provide better electrical characteristics for neural network applications.

We have also incorporated the SONOS synaptic weights into a single-level two tap linear adaptive neuron employing a Widrow-Hoff's delta learning rule. The combination of CMOS control circuits and SONOS synaptic weights has demonstrated the feasibility of integrating these two technologies onto a single silicon wafer. The initial results are encouraging and promising and

provide insight and direction into the integration of these two technologies to realize large artificial neural network systems.

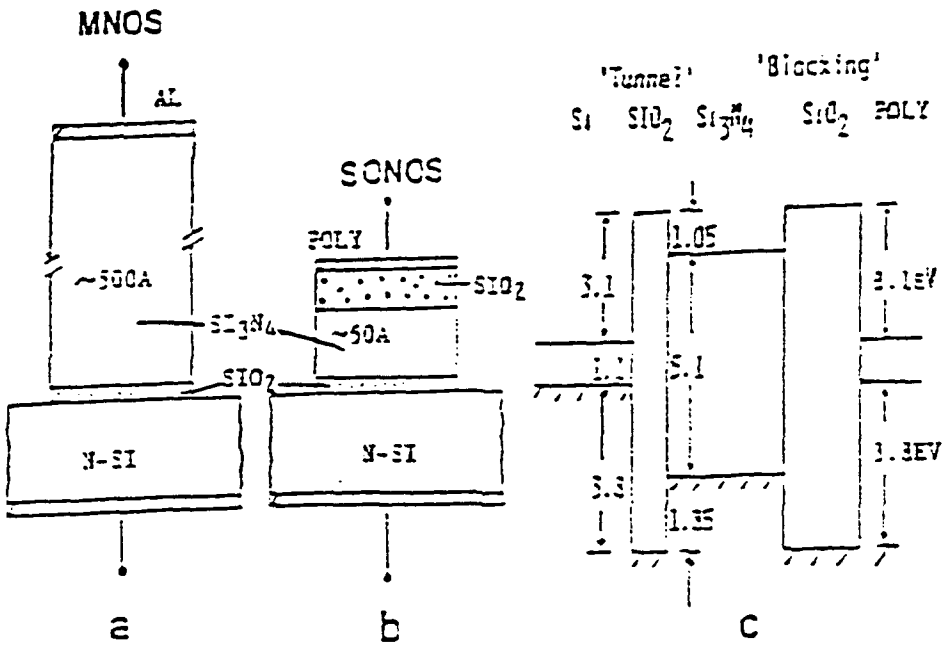


Figure 1. Cross Sectional View of the (a) MNOS (b) SONOS Electrically Modifiable Synaptic Weight (c) SONOS Ideal Energy-Band Diagram

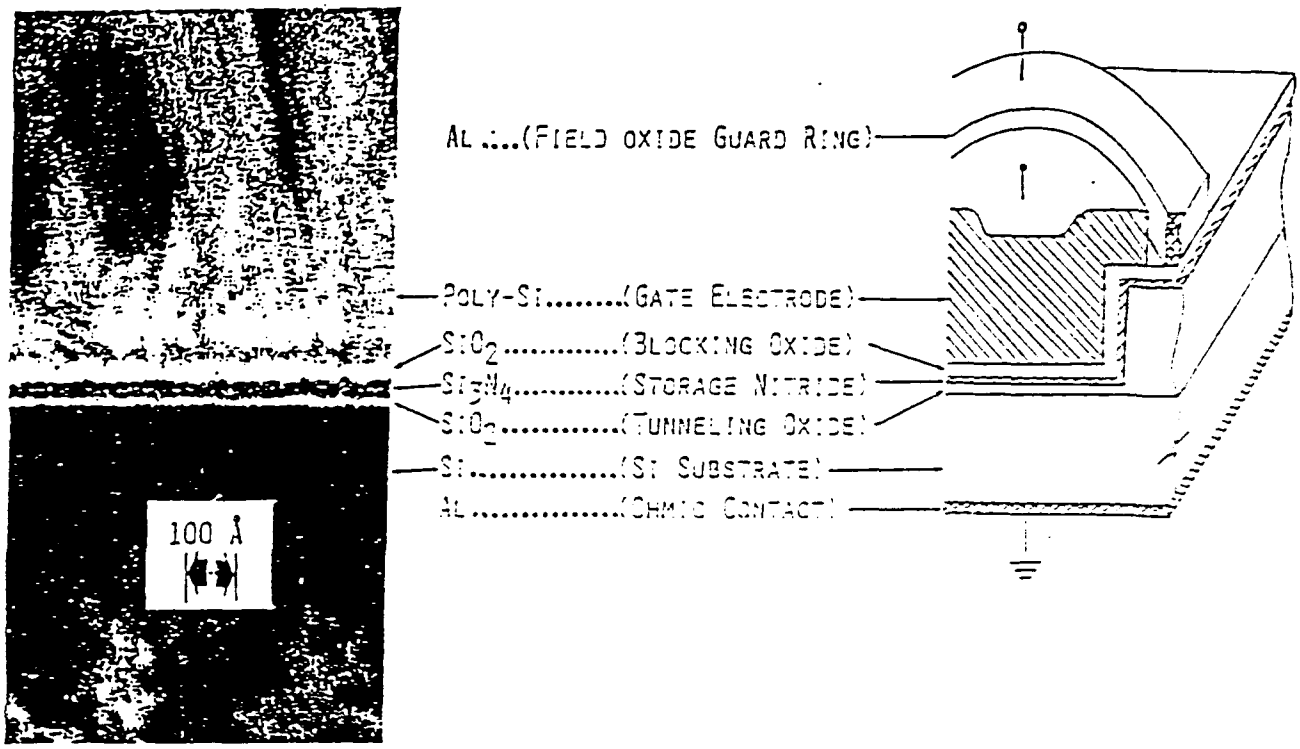


Figure 2. TEM Photomicrograph of the SONOS Synaptic Weight

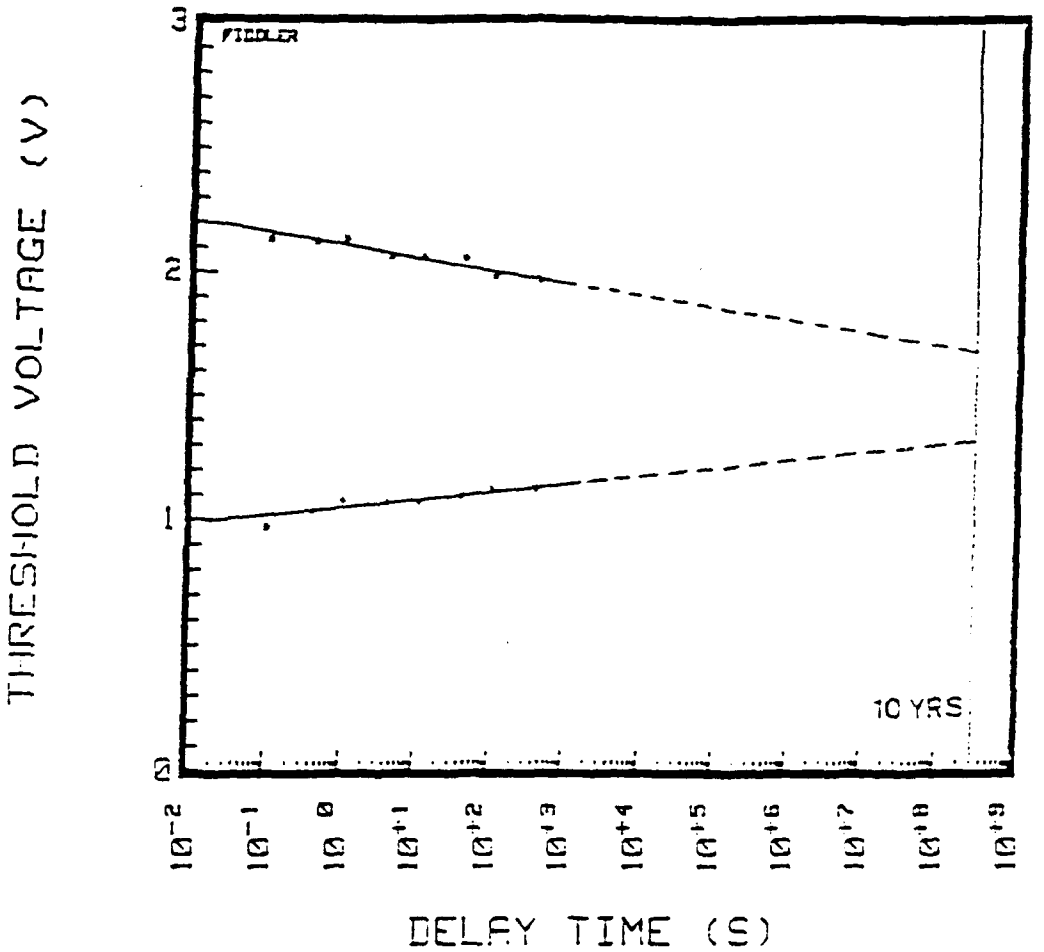


Figure 3. Retention Characteristics of a Modifiable SONOS Synaptic Weight

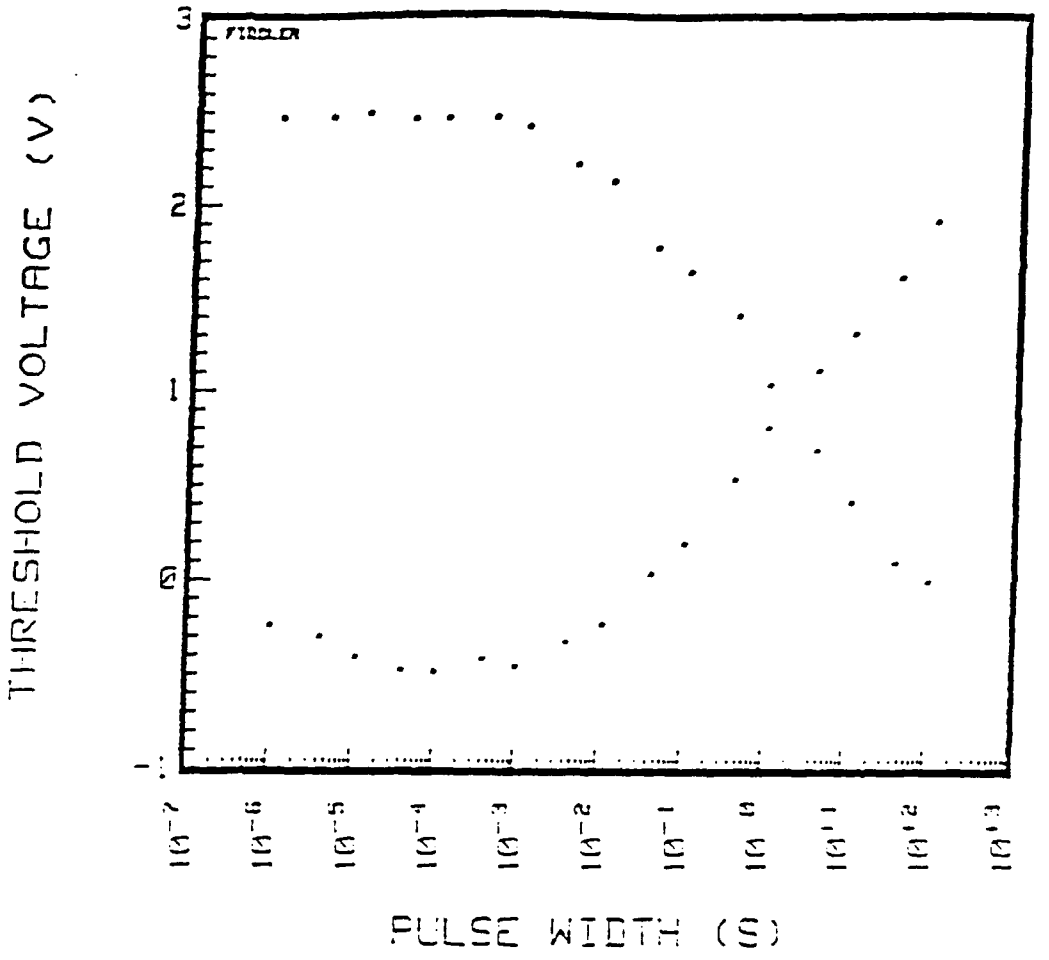


Figure 4. Erase/Write Characteristics of a Modifiable SONOS Synaptic Weight

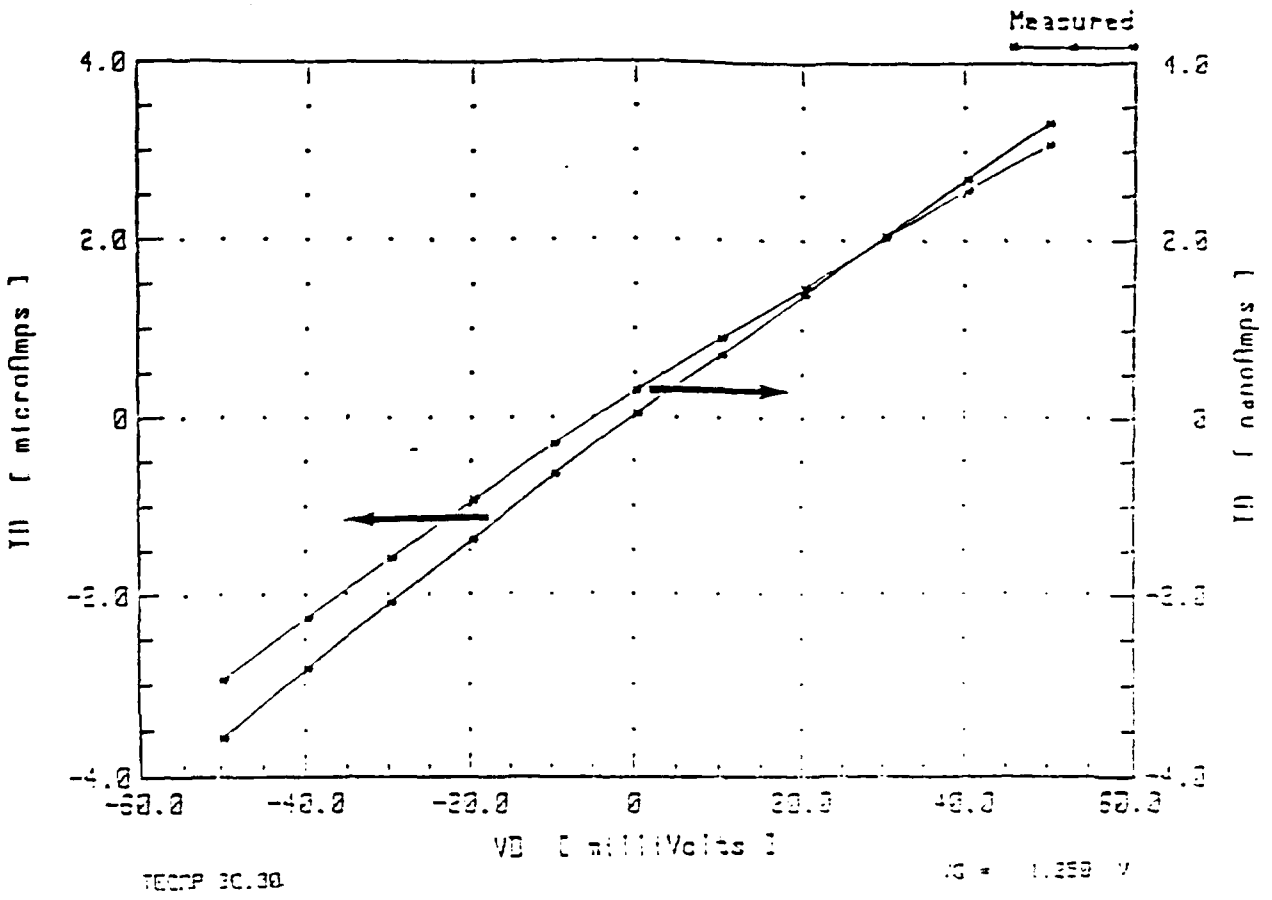


Figure 5. Dynamic Range of a SONOS Synaptic Weight After Programming

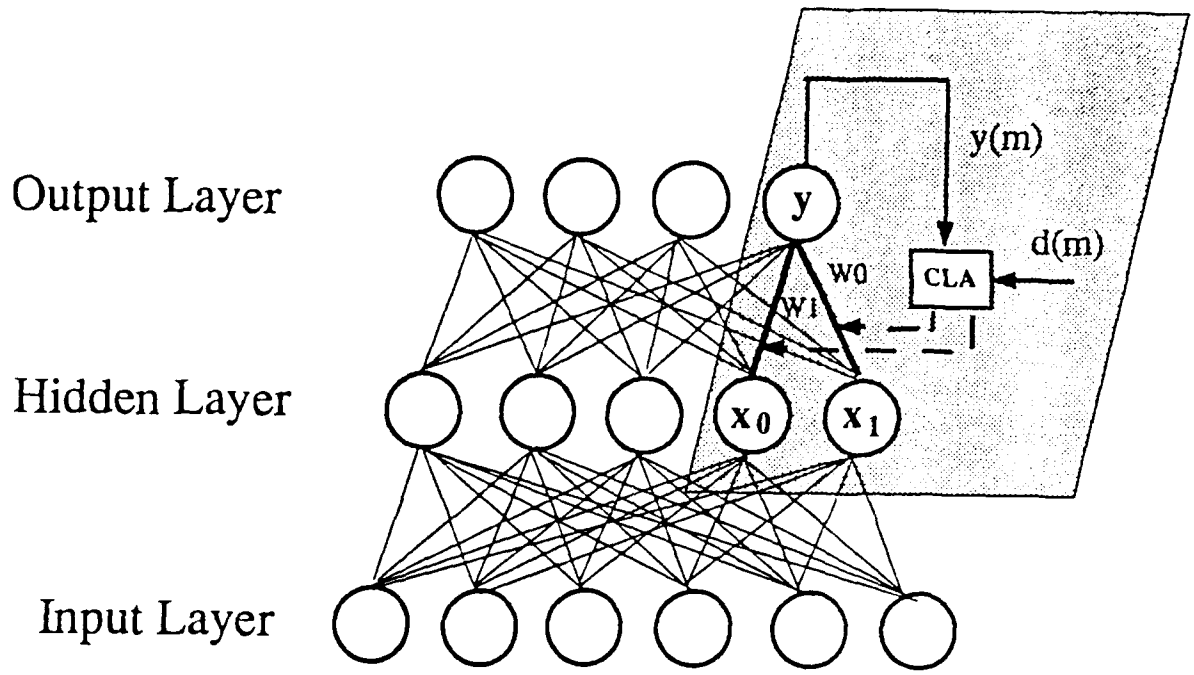


Figure 6. Conceptual View of Multi-layer Artificial Neural Network Architecture Incorporating the Single-Level Linear Adaptive Neuron

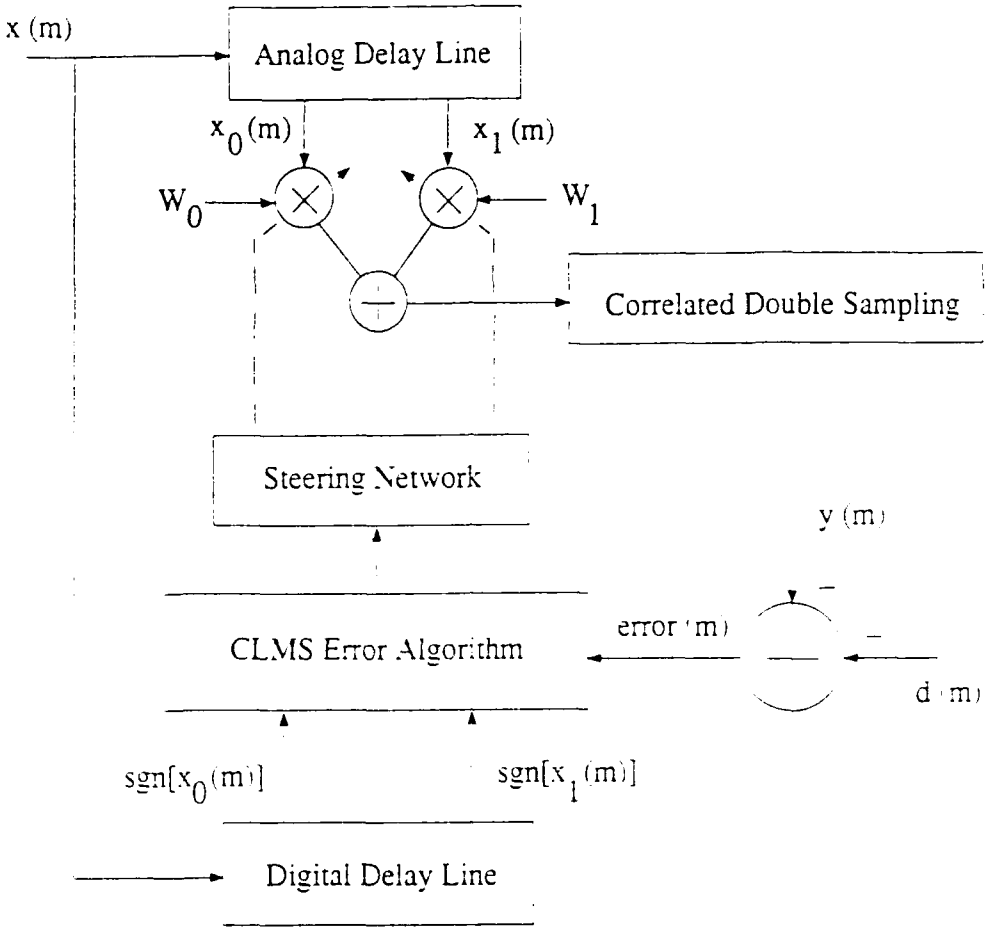
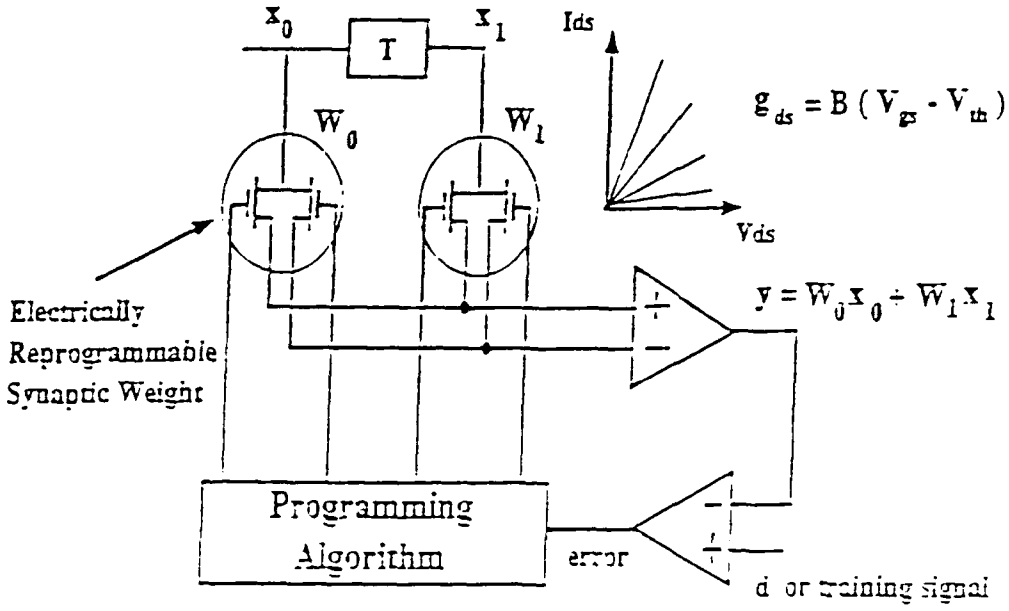


Figure 7. Block Diagram of a Single-level Linear Neuron



$$error = d - y$$

$$W_j = C (g_{is0}^{+} - g_{is0}^{-})$$

$W_j > 0$ Excitatory Synaptic Weight

< 0 Inhibitory Synaptic Weight

Figure 8. Electrical Implementation of the Synaptic Weights

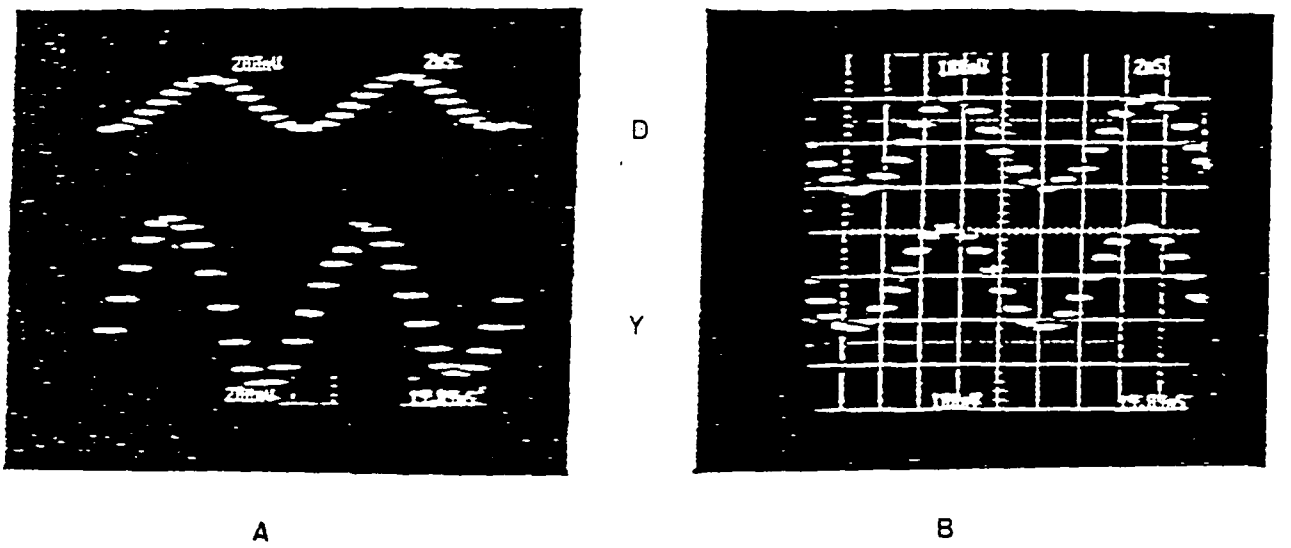
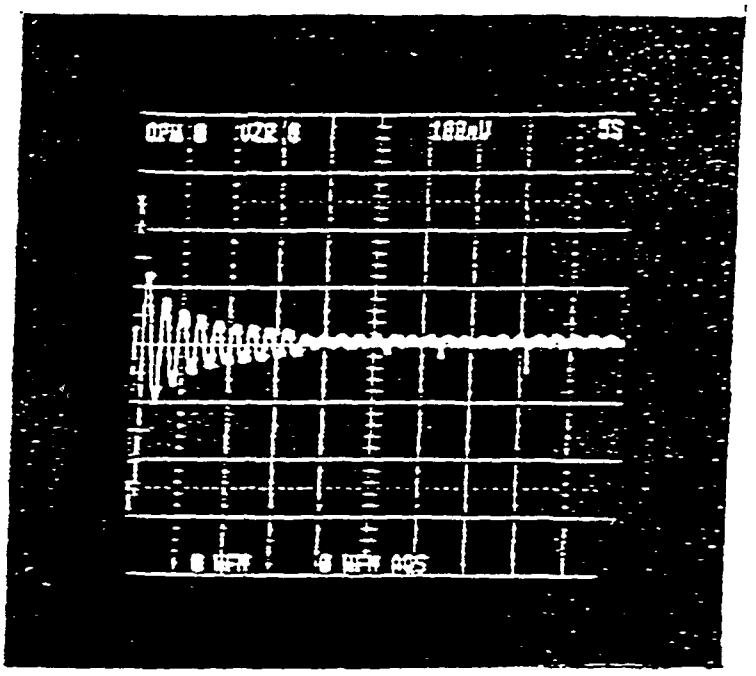


Figure 9. Output and Training Signals versus Time Characteristics (a) Initialized (b) Adapted

ERROR



TIME

Figure 10. Error Signal versus Time Characteristics

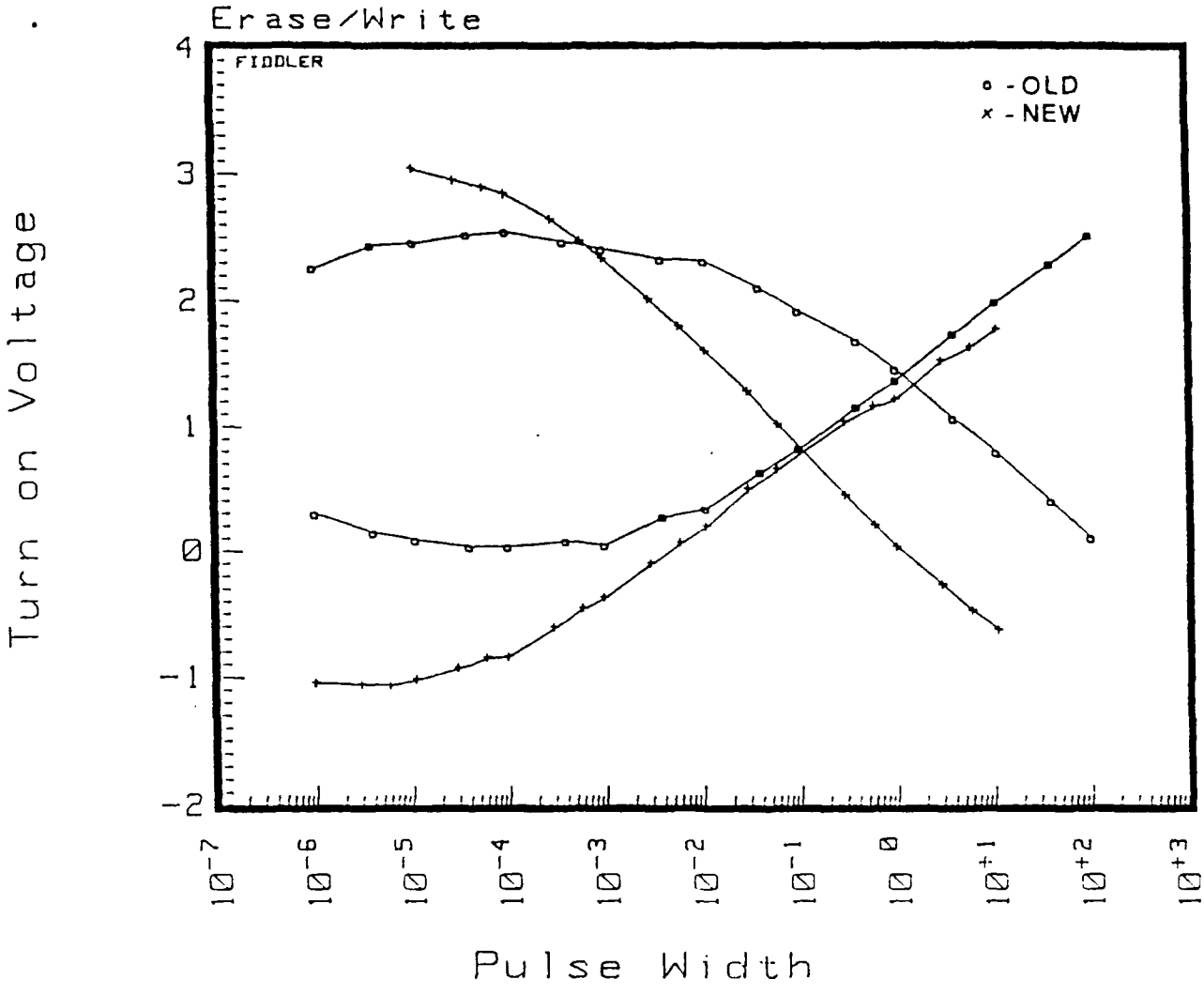


Figure 11. Improved Characteristics Demonstrated by the Newly Fabricated SONOS Synaptic Weight Elements

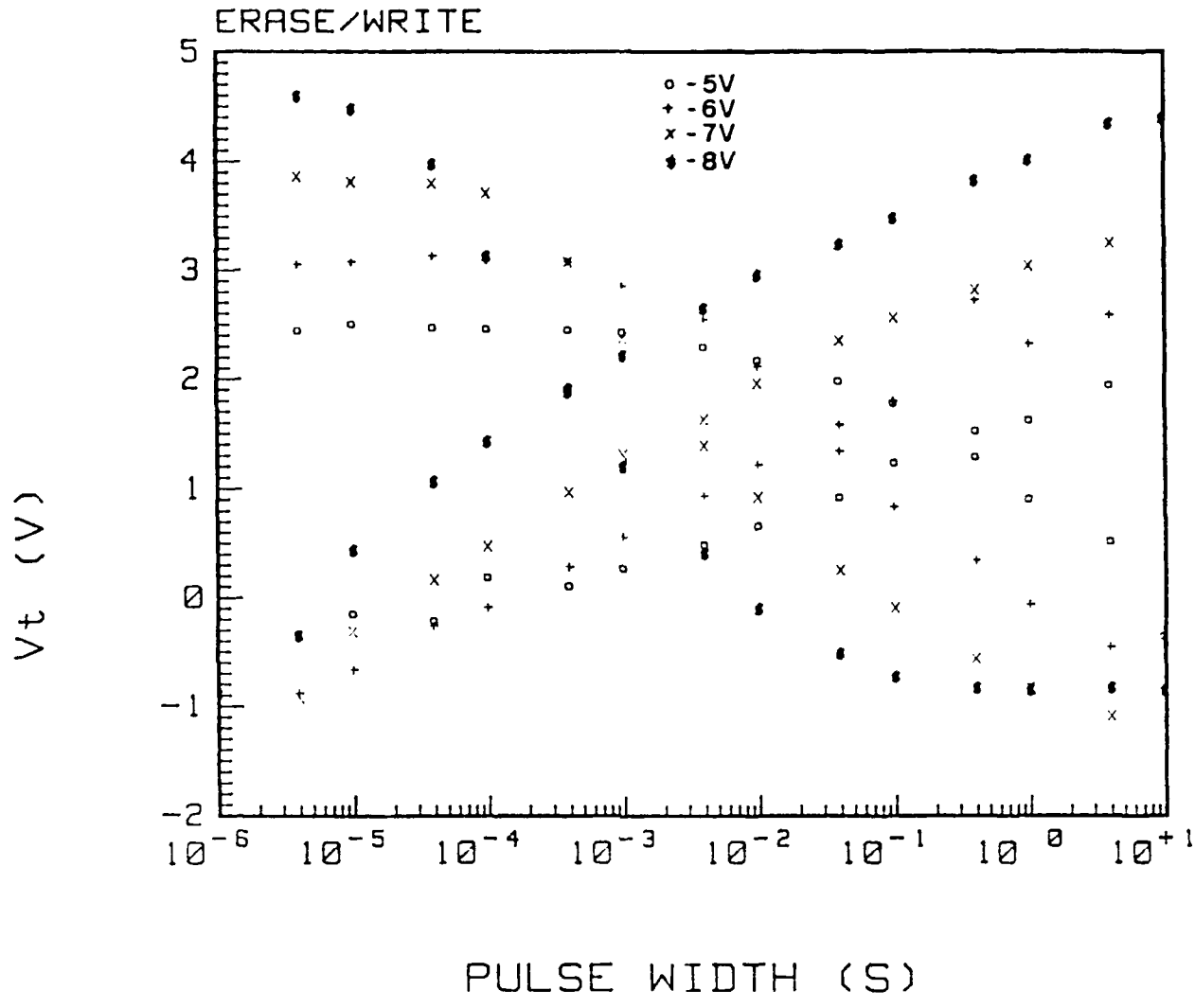


Figure 12. Programming Voltage Dependence in the Programming Speed

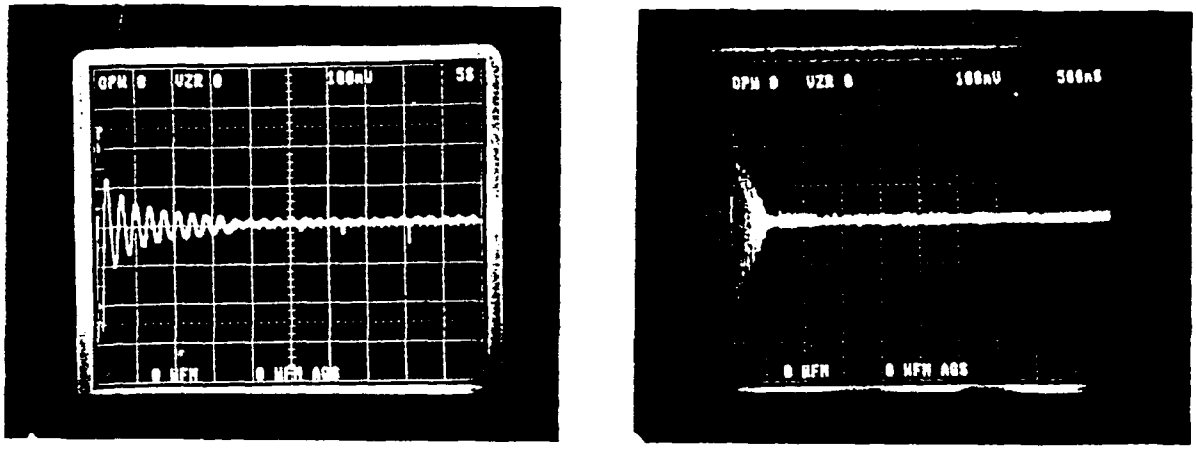


Figure 13. Comparison between Old and New Convergence Time Performance

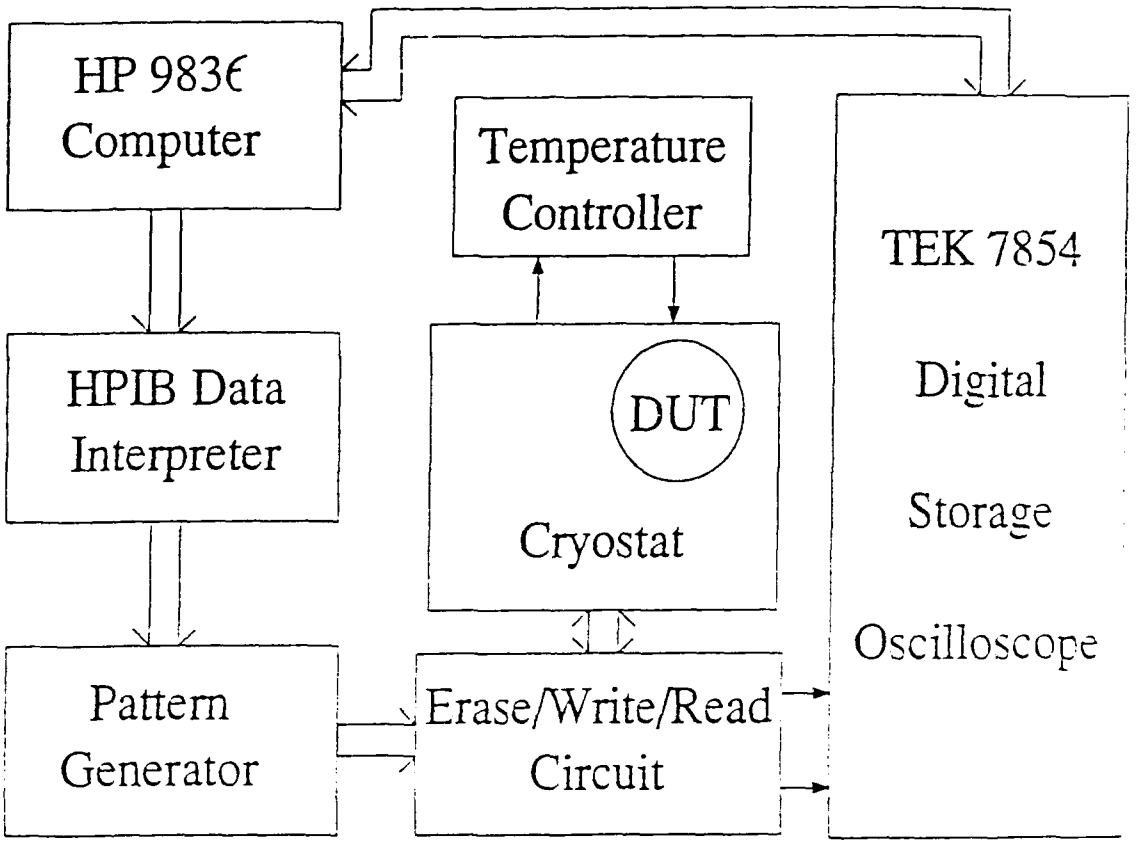


Figure 14. Block Diagram of the Automated Data Acquisition System

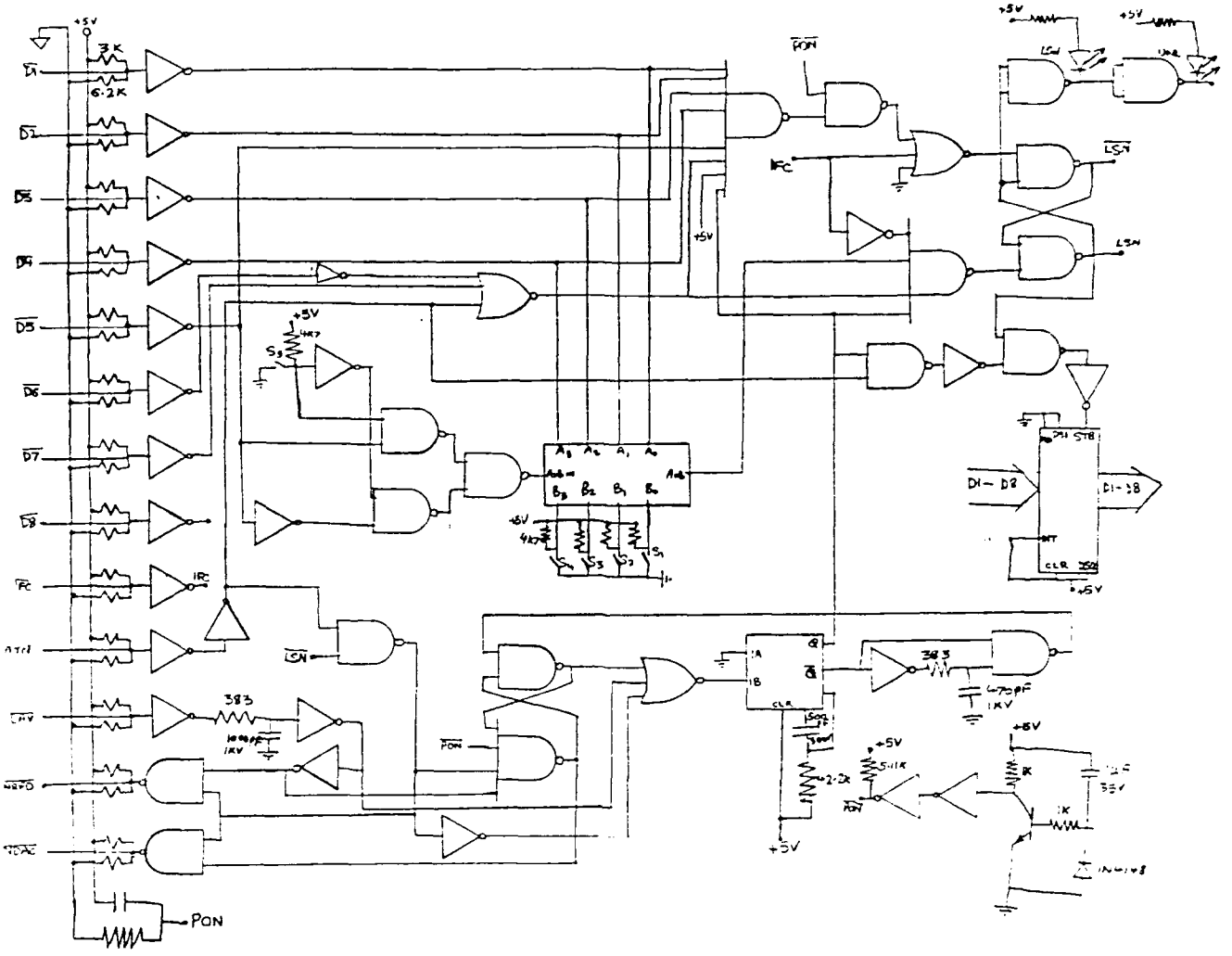


Figure 15. Circuit Schematics of the HPIB Command/Data Interpretor

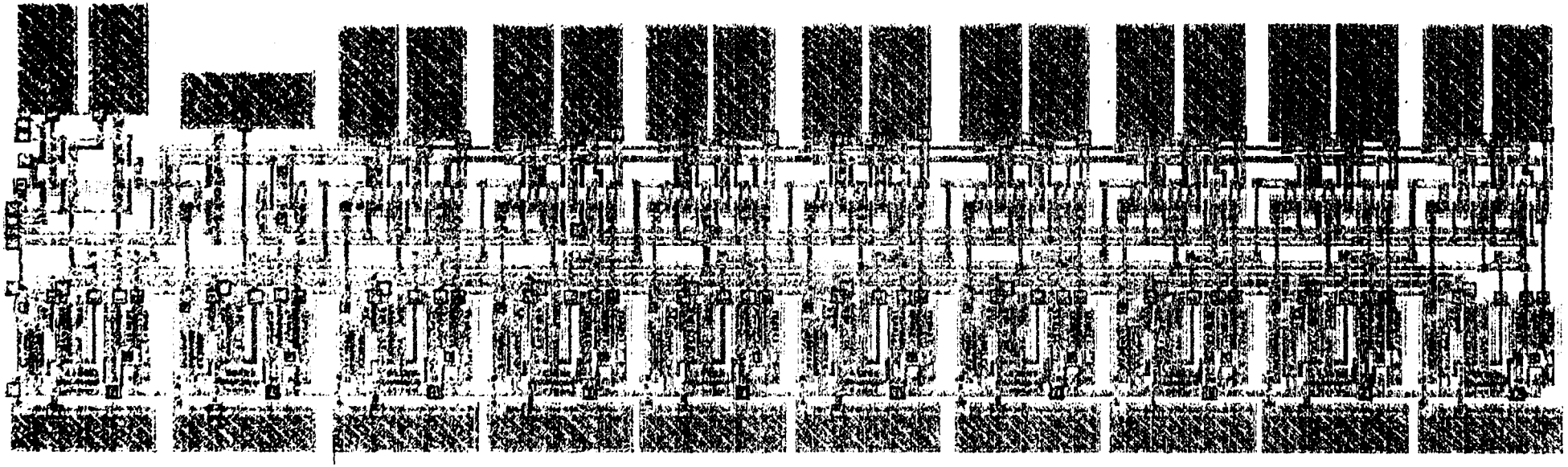


Figure 16. Sample Layout of the Analog Delay Line

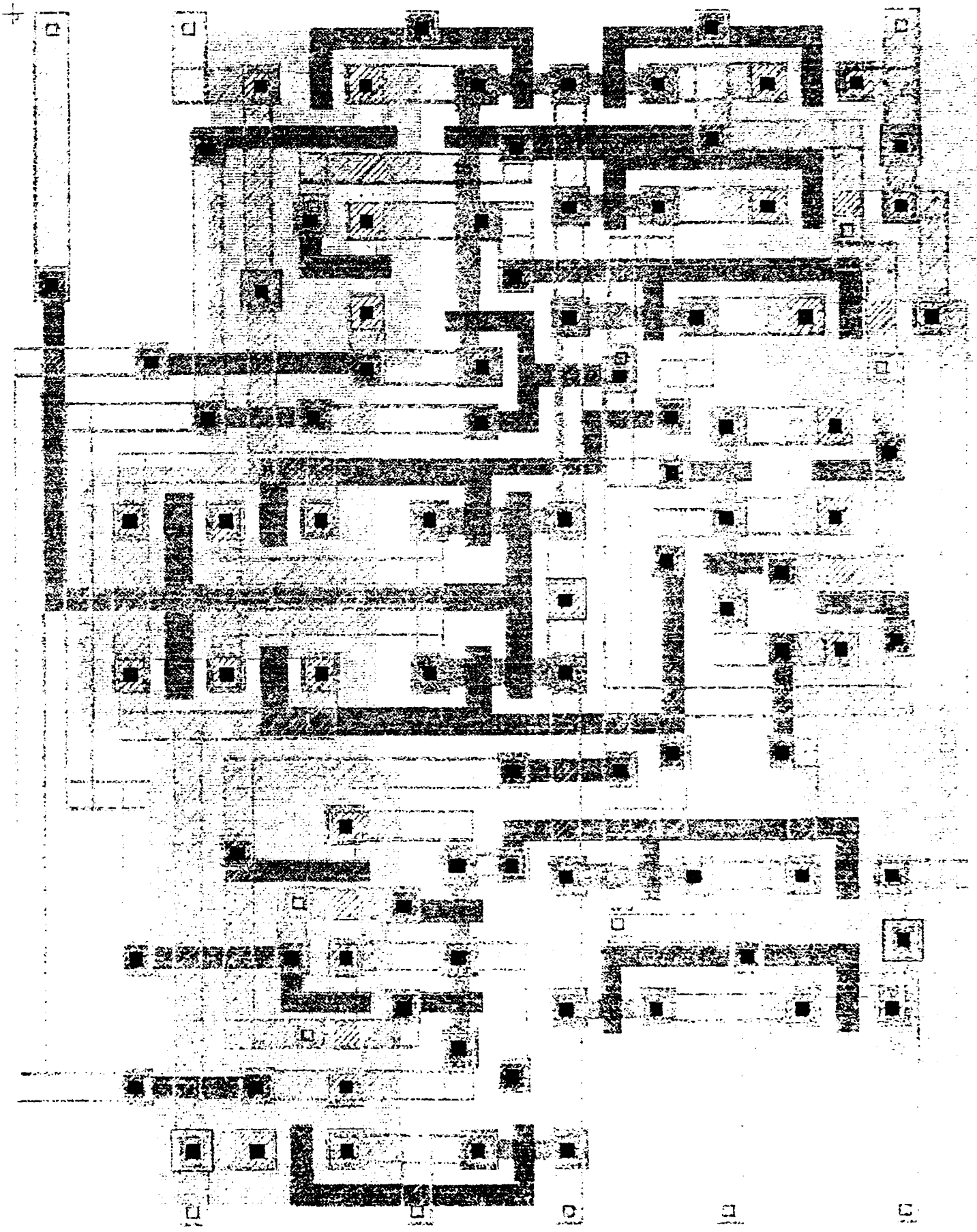


Figure 17. Sample Layout of the D Flip Flop

References

1. D.O. Hebb, *The Organization of Behavior*, John Wiley, 1949.

2. T. Kohonen, *Associative Memory: A System Theoretic Approach*, Springer-Verlag, 1977.

3. D. Rumelhart, and J. McClelland, *Parallel Distributed Processing*, MIT Press, 1986.

4. John J. Hopfield and David W. Tank, "Computing with Neural Circuits: A Model", *Science*, Vol. 233, August 1986, pp. 625-633.

5. J. Denker, "Neural Network for Computing: Snowbird 1986", *AIP Conference Proceedings*, 1986.

6. H.P. Graf and P. deVegvar, "A CMOS Implementation of a Neural Network Model", *Proceedings of Stanford Conference on Advanced Research in VLSI*, 1987.

7. Y. Tsividis and S. Satyanarayana, "Analog Circuits for Variable-Synapse Electronic Neural Networks", *Electronic Letters*, Vol. 23, No. 24, Nov 1987, pp. 1313-1314.

8. F.J. Kub, I.A. Mack, K.K. Moon, C.T. Yao and J.A. Modla, "Programmable Analog Synapses for Microelectronic Neural Networks Using a Hybrid Digital-Analog Approach", *IEEE Device Research Coference on Neural Networks*, 1988, pp. 24-27.

9. Mark Holler, Simon Tam, Hernan Castro, and Ronald Benson, "An Electrically Trainable Artificial Neural Network (ETANN) with 10240 'Floating Gate' synapses", *Proceedings of IJCNN*, 1989.

10. F.R. Libsch, A. Roy and M.H. White, "A True 5V EEPROM Cell for High Density NVSM", *IEEE Trans. Electron Devices*, Vol. ED-34, No. 11, 1987, pp. 2371.

11. M. White, D. Lampe, F. Blaha and I. Mack, "CCD and MNOS Devices for Programmable Analog Signal Processing and Digital Nonvolatile Memory", *IEEE International Electron Devices Meeting*, 1973.

12. M. White, D. Lampe, F. Kub, and D. Barth, "A Nonvolatile Charge Addressed Memory (NOVCAM) Cell", *IEEE. J. of Solid-State Circuits*, October 1975.

13. M.H. White, I.A. Mack, G.M. Borsuk, D.R.Lampe, and F.J. Kub, "Charge-Coupled Device (CCD) Adaptive Discrete Analog Signal Processing", *IEEE J. of Solid-State Circuits*, Vol. SC-14, 1979, pp. 132.

14. J. Sage and R. Withers and K. Thompson, "MNOS/CCD Circuits for Neural Network Implementation", *IEEE International Symposium on Circuits and Systems*, 1989, pp. 1207-1209.

15. B. Widrow and M. Hoff,Jr., "Adaptive Switching Circuits", *IRE WESCON Conv. Rec.,pt. 4*, 1960, pp. 96.

16. M. White and I. Mack, *A CCD Monolithic LMS Adaptive Analog Signal Processor Integrated Circuit*, Final Report, Contract N00173-77-C-0328, 1980.

17. M. White, "A VLSI Conductance Multiplier (Synapse) for Hardware Implementation of Neural Networks", *invited paper to workshop on Hardware Implementation of Neural Networks*, 1988.

18. Chun-Yu Malcolm Chen, Margaret L. French and Marvin H. White, "An Analog Nonvolatile Eletrically Modifiable Synaptic Element for VLSI Neural Network Implementation", *11th IEEE Nonvolatile Semiconductor Memory Workshop*, 1991.

19. Anirban Roy, "Retention, Endurance and Interface Traps in MONOS Memory Transistors", Master's thesis, Lehigh Univeristy, 1985.

20. T.J. Krutsick, M.H. White, H.-S. Wong and R.V.H. Booth, "An Improved Method of MOSFET

Modeling and Parameter Extraction", *IEEE Trans. on Electron Devices*, Vol. ED-34, No. 8, Aug. 1987, pp. 1676-1680.

21. Umesh Sharma and Marvin H. White, "Ionization Radiation Induced Degradation of MOSFET Channel Frequency Response", *IEEE Trans. on Nuclear Science*, Vol. 36, No. 3, June 1989, pp. 1359-1366.

22. Marvin H. White, Donald R. Lampe, Franklyn C. Blaha, and Ingham A. Mack, "Characterization of Surface Channel CCD Image Array at Low Light Levels", *IEEE J. of Solid-State Circuits*, Vol. SC-9, No. 1, February 1974, pp. 1-13.

23. J.L. Moschner, "Adaptive Filter with Clipped Input Data", Tech. report, Standford Lab. Report, No. 6796-1, June 1970.

24. Marvin H. White and Chun-Yu Chen, "Electrically Modifiable Nonvolatile Synapses for Neural Networks", *Proceedings of IEEE International Symposium on Circuits and Systems*, 1989, pp. 1213-1216.